The Behavior Analyst 2012, 35, 29–35 No. 1 (Spring)

The Central Role for Behavior Analysis in Modern Robotics, and Vice Versa

William R. Hutchison
Behavior Systems LLC

Serving as Rachlin's teaching assistant for his graduate course on animal learning in 1973 determined the direction of my career, which has been to build computer models and robots based mostly on the equations from that course and related ones. These artificial beings behave and learn very much like animals, and creating them forces a far more molecular perspective on behavior than Rachlin takes in addressing the issues in this article. These comments will address Rachlin's points from the other perspective to which he has contributed enormously, the quantitative analysis of behavior.

Selectionist concepts, which as Rachlin (2012) points out encompass both learning and biological evolution, have become very important in robotics and computer intelligence. This direction has been driven by the practical realization that many things that we want computers and robots to do are so complex as to be infeasible to analyze and program, but can be learned by computer programs that emulate these biological processes. One important technology, actually called *reinforcement learning* (Sutton & Barto, 1998), was originally based on one of the very equations taught in Rachlin's course, by Rescorla and Wagner (1972). The approach by Edelman (1987) that Rachlin describes is a behavioral model nominally based on brain function, but was obviously shaped by his Nobel Prize-winning breakthroughs applying evolutionary concepts to the immune system and constrained by the need to produce operant learning. Among behavior analysts, Donahoe and Palmer (1994) have developed a sophisticated model that proposes how operant behavior is implemented in the brain, informed by detailed knowledge of operant learning. My own operant computer models and applications, based on operant learning, were independently developed starting at the same time as those early models (e.g., Bell, Hutchison, & Stephens, 1992; Forsyth, Hawkins, & Hutchison, 1996; Hutchison, 1996, 1997, 1998; Hutchison & Constantine, 2003, 2005; Hutchison, Constantine, Borenstein, & Pratt, 2007, 2008a, 2008b; Hutchison & Stephens, 1987; Stephens & Hutchison, 1992, 1993; Stephens, Hutchison, Hormby, & Bell, 1990). Many others have adopted learning methods for robotics (Dorigo & Colombetti, 1998; Lin, 1992; Touretzky & Saksida, 1997; Weng & Chen, 1996).

It is quite telling that all of these systems assume the basic formulations of operant conditioning: that certain biologically significant events (i.e., primary reinforcers and punishers) function to select (i.e., increase or decrease) the strength of relations between recently occurring stimuli and recently emitted behaviors (i.e., stimulus control). In addition, all of these systems incorporate the key process of higher order conditioning, in which the stimulus control values acquired during primary reinforcement provide the basis for conditioned reinforcement. Although everyone recognizes that these relations

Correspondence concerning this paper should be addressed to William R. Hutchison, Behavior Systems LLC, 5475 Tenino Avenue, Boulder, Colorado 80303 (e-mail: whutchi@behaviorsys.com).

are implemented by the brain, the models can consist of purely molecular behavioral relations (e.g., Rescorla & Wagner, 1972) expressed in equations, not claiming point-to-point correspondence with neural processes. However, the author acknowledges that trying to integrate behavioral and neural models also has much potential value, because each provides information that may guide the other, especially in the direction of neural models needing to function according to known behavioral processes. It is frustrating that most practitioners, in both robotics and neuroscience, have little or no awareness of the extensive knowledge of operant learning.

At about the same time the foregoing operant learning methods were developed, another major selectionist approach was invented to emulate the evolutionary processes of variation, selection, and retention and reproduction, known as genetic algorithms or GAs (Holland, 1985). Rachlin points out that both learning and evolution are selectionist, but they are not interchangeable or competitive. They collaborate in an obvious way. Evolution determines the fixed characteristics of animals that persist during lifetimes to determine the organismic elements that enter into learning during that lifetime: the characteristics of our sensors, the kinds and characteristics of our actions, and the characteristics of the primary values and learning process that enable acquisition of intelligent control of actions. Many practitioners have applied GAs to evolve static behavior control that does not change from experience, but a far more exciting idea is to implement the interplay of learning and evolution as in nature. This works by creating populations of heterogeneous operant beings, each of which learns from experience and reproduces differentially to produce the next generation based on their success in the given environment. The environment can be the real world, but by implementing it in computer virtual reality, the entire process can play out millions of times faster than evolution occurs in the real world. The speed and power of such a robot evolution and design system, and the fact that it operates with no human intervention, have ominous implications for robots as a competitive species (discussed below).

Using the computer as a medium to unambiguously express and test sets of posited quantitative relations has been essential to most sciences, and it is equally valuable in behavior analysis: It enables us to express all our formulations and their interrelations in a cohesive way and then project their implications in a disciplined way that would be impossible with separate equations. The first impact of such modeling is on ourselves, to discover that some of our assumptions do not play out in practice exactly as we assumed. That is, they do not behave like our animal subjects would under the same circumstances. Simulations in these models facilitate adjusting our formulations to ones that produce the behavior we predict, a strategy sometimes called analysis by synthesis. For example, the delay-discounting (DD) functions that Rachlin describes are often built into these operant computer models, but testing with the models shows that even with DD functions the direct effect of reinforcement can only be effective for delays in the range of seconds, because the number of behavioral relations that might be reinforced by a delayed consequence grows too large with time; this is known as the credit assignment problem. We can still model the development of molar DD functions over more time, but we see them produced by conditioned reinforcers and, over much longer periods such as months, by verbal mediation. This is fertile ground for the molar versus molecular debate, as Rachlin's response will hopefully enlighten, but this kind of computer modeling clearly forces its practitioners toward a molecular perspective that must produce observed molar relations.

The second major impact of these models, once they work, is that they enable us to prove that our posited behavioral processes are sufficient to produce the behavioral phenomena of interest. Demonstrations using a computer model function logically as sufficiency proofs that refute claims, often suffered by behavior analysts, that our relatively simple concepts are inadequate to explain complex behavior. These operant models have demonstrated (reported in the Hutchison references below) almost all of the fundamental behavioral processes including generalization, learning abstract visual discriminations, multiple coordinated movements, higher order conditioned reinforcement, chaining, response variability that changes as with animals, stimulus equivalence phenomena, and Skinner's (1957) verbal operants, including minimal echoics (enabling phoneme imitation needed for humanlike language teaching), tacts (nouns, verbs, adjectives), intraverbals (facts and rules), self-echoics, compliance with commands, autoclitics (generalized adjective–verb and subject–verb word order), and true rule-guided behavior (tacting a situation, which evokes a relevant intraverbal rule, then performing the behavior specified in the rule). Such demonstrations of operant computer models are most persuasive when they interact directly with the real world, like simple robots, rather than only inside a computer, not just because they are most visible but because they remove all the ways that we could be cheating, both intentional and unintentional. A real Turing test such as Rachlin discusses is performed every year, known as the Loebner Prize, formerly conducted under the auspices of the Cambridge Center for Behavioral Studies. The current author twice served as the technical director of that contest, and the clever tactics of the computer programs that win the contest are enlightening but are far from true verbal behavior produced in an operant system. As Rachlin points out, they appear human under constrained circumstances, but interacting with them under a wider range of conditions inevitably exposes the programmed control behind them.

The third impact of these models of operant behavior is that when they are embodied in robots, they produce synthetic animals that can perform valuable functions in the real world. As Rachlin points out, the economic benefits of these robots function to reinforce behaviors of providing resources to produce and further develop the robots; a case of natural evolution even though not based on carbon DNA. Contrary to popular depictions, most intelligent robots will not be humanoid in form. Humans are generalists whose bodies and brains enable performing a huge variety of behaviors in many environments, but robots developed to perform a limited range of tasks will have an extremely wide range of bodies optimized for their tasks and only those behaviors and knowledge needed for those tasks. Robots that provide care or companionship for humans should probably be more humanoid, such as the remarkably lifelike Actroid robots developed by the Kokoro Company with Osaka University (Christensen, 2005).

CAN ROBOTS BECOME HUMAN OR SUPERHUMAN?

Rachlin uses the watershed event of a computer surpassing humans in an impressive ability in order to create a thought exercise that forces readers to reflect on the nature of humans. This exercise is valuable, even though it will persuade few people to define human simply as any being that behaves identically to humans. Behaviorists are perhaps more likely than most to agree with Rachlin that humans can nevertheless form strong (loving?) relationships with sophisticated robots due to the positive value they can provide, even knowing him or her (following Rachlin's usage) to be a robot. The movie that most realistically portrays how behavioral theory would project the future of adaptive robot technology is *Bicentennial Man*, based on Asimov's science fiction novel (1996). A factory-pretrained robot with learning abilities becomes a servant in a human household and adapts to their family life. Even though everyone knows "he" is a robot, the movie shows realistically how he naturally acquires behaviors from interacting with humans and the world that make him quite endearing and humanlike. The movie plays out Rachlin's thought exercise effectively, and we could press the argument further by asking whether the viewer would consider him to have human rights (e.g., not being allowed to be killed). David Levy's (2007) book *Love and Sex with Robots* provides a survey of the history of human companionship and sexual practices along with current and envisioned technology. He makes a surprisingly strong case that humans will accept sophisticated robots as sexual partners and even primary companions, largely because of their ability to provide many reinforcers accompanied by few punishers.

Behavior analysts are the best suited to lead the development of intelligence in these emerging robots, and doing so would be valuable for us. The technology gives us better ways to express and test our scientific formulations, it proves to outsiders the sufficiency of our formulations, and it gives us valuable roles in an important industry, in which we can provide extremely relevant expertise of which those outside behavior analysis are largely ignorant.

However, working with these robots and projecting the future does raise some very important concerns. For example, endowing robots with a complete operant system including their own primary values, such as electric power (which as Rachlin points out will be different from humans) has a huge practical value in enabling robots to learn from their own experience rather than needing the constant delivery of consequences by a human teacher. However, there is an unanticipated consequence of doing that: The behavior of such robots is selected only by those values. People might expect a person's own robot to always follow commands from its owner, and with proper primary values and training it can learn to comply just like humans do. However, one can easily recognize that such a robot will disobey the human under the same circumstances in which a human would. A robot can have collision sensors and primary values that included collisions as punishing events. The author trained such a robot to comply with commands to move, but when commanded repeatedly to ram into a wall, it eventually "refused" because the punishment weakened the compliance behavior in that situation. Rachlin discusses whether such a robot felt "pain." Certainly it felt the sensation that had the punishing effect, so "sort of." What is still lacking to more clearly label it "feeling pain" is that the robot has not learned, as humans do, to tact (i.e., emit a verbal label in response to) those events as pain, in order to be aware or conscious of the punishing event. But it could easily learn to do so: Rachlin points out that tacting the event to others (either crying "ouch!" or, Spock-like, calmly reporting "damaging event") can evoke positive responses from others, and for that reason might be exaggerated, as could easily be learned by our robot. Stretching into the moral sphere as Rachlin discusses, the robot could learn to tact (be aware) that the situation contains a choice whether to be deceptive and to think (self-talk) about the consequences before doing it or not. The key point here is that an operant robot is an independent being, even though he is dependent on others, just as all humans are dependent on others.

Rachlin projects a future Watson II that is sophisticated and humanlike. However, we behavior analysts have a vital interest in extending our own analysis further, because it suggests that the inexorable evolutionary process that plays out both in the real world and in accelerated virtual reality is destined to produce robots that will displace humans at the top of the pyramid, just as all previous species have been displaced. Such an outcome seems like the ultimate folly of humans, to produce a species that takes control from them. Most people considering this prospect believe they have a solution to prevent this, but our own behavioral analysis leads to rejecting all the proposed solutions. Arguments against robots taking over, and the flaws in the arguments, include the following:

1. Computers cannot replicate some mechanisms that humans have. But as Rachlin points out, we assume that humans are physical beings whose mechanisms consist of physical structures. If the mechanism's function can be described, computer science should be able to replicate the function even if not by the same physical mechanism. That is exactly what we have seen in producing robots with operant learning and biological evolution.

2. Human brains with billions of neurons and trillions of connections have more power than can be produced by a computer. But Moore's law (1965), which has accurately predicted the exponential growth of computing power, tends to undermine this argument, and many "serial" human behaviors such as numerical and logical computation have already been dwarfed by computers. Parallel computer architectures and futuristic computational methods such as optical and quantum computing promise computing power vastly beyond that of human brains.

3. Humans cannot design a robot that is more intelligent than themselves. Watson is an example that supports this argument, because the developers collected human knowledge and programmed it in a computer, so it barely exceeds an individual human's ability and cannot go much further. However, the operant and evolutionary technologies discussed here do not depend on human participation, but instead involve a hyperspeed evolutionary process that produces increasingly advanced beings that can learn on their own. I have consistently found that the designs produced by a combined learning and evolutionary process are superior to my own designs for most robotic problems. Beyond rapid learning of single robots, the physical structure and learning of a computer system can easily be cloned, so new robots will not have to go through the massive learning process required of humans, but can begin "life" with Lamarckian-like feature inheritance plus knowledge from the most advanced existing robot from which they can continue learning. Note also that robots are already manufacturing computers and the robot bodies increasingly independently.

4. Humans should be protected by Asimov's (1942) rules for robots; robots must not harm humans and must obey human orders. However, Asimov was assuming robots would be programmed, not based on operant systems. Behavior analysts recognize the nature of rules in operant systems: They are learned, not hard wired in brains. Even if it were possible to mandate that all robots be trained to obey these rules, humans have amply demonstrated the obvious: Rules can be disobeyed.

5. Robots should be programmed with primary values to promote human well-being, which is a more

behavioral strategy consistent with Rachlin's analysis. However, as Skinner has pointed out, primary values are themselves evolved to select behavior that is fittest in an environment. For example, animals not reinforced by food would not eat and animals not reinforced by sex would not mate. But even if it were mandated to hard wire certain positive values, there is no way to enforce the mandate universally, and subsequently they could be changed accidentally or on purpose by hackers, who constantly demonstrate their creative destructiveness. A robot could even potentially access its own or another robot's control program, unlike humans, and could change primary values accidentally or intentionally.

6. Finally, as in the movie *2001 Space Odyssey*, in which, as Rachlin points out, flaws in the computer's intelligence make it lethal, many assume that humans could ultimately just pull the plug if robots threaten human superiority, or more proactively, prohibit research and development of robots that would be capable of doing so. In this multilateral political world, there is no way to enforce this (witness bans on research areas such as cloning) apart from intentional hacking. And we are already nearly at the point at which the development and manufacturing of such robots need no human participation.

SUMMARY

Rachlin has demonstrated how the existence of intelligent robots might stimulate thoughtful humans to adopt more effective understanding of our own nature. Behavior analysts also have much to gain from participating in the development of computer and robotic models of operant behavior: We can express and test our formulations more effectively, we can prove the sufficiency of our formulations to others, and we can play important roles in the development of valuable applied technology. The technology seems destined to threaten human well-being in the future, but that prospect should encourage more rather than less participation by behavior analysts.

REFERENCES

Asimov, I. (1942, March). Runaround. *Astounding Science Fiction*, 94–103.

Asimov, I. (1996). *The bicentennial man*. New York: Grafton.

Bell, T. M., Hutchison, W. R., & Stephens, K. R. (1992). Using adaptive networks for resource allocation in changing environments. In P. J. G. Lisboa (Ed.), *Current perspectives in neural networks* (pp. 180–190). London: Chapman and Hall.

Christensen, B. (2005, June 28). New robot looks strikingly human. *LiveScience*. Retrieved from http://www.livescience.com/

Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston: Allyn and Bacon.

Dorigo, M., & Colombetti, M. (1998). *Robot shaping: An experiment in behavior engineering*. Cambridge, MA: MIT Press.

Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.

Forsyth, J. P., Hawkins, R. P., & Hutchison, W. R. (1996). Neural network learning theory: Can it help resolve the behavioral cognitive controversy? *The Behavior Therapist*, *19*, 5–9.

Holland, J. H. (1985). Properties of the bucket brigade algorithm. In J. J. Grefenstette (Ed.), *Proceedings of the first international conference on genetic algorithms and their applications* (pp. 1–7). Mahwah, NJ: Erlbaum.

Hutchison, W. R. (1996). We also need complete behavioral models. *Journal of the Experimental Analysis of Behavior*, *67*, 224–228.

Hutchison, W. R. (1997). *Learned emergence of functional symbol systems in adaptive autonomous agents* (pp. 287–292). Proceedings of the International Conference on Intelligent Systems and Semiotics, Gaithersburg, MD, NIST.

Hutchison, W. R. (1998). Computer simulations of verbal behavior for research and persuasion. *The Analysis of Verbal Behavior*, *15*, 117–120.

Hutchison, W. R., & Constantine, B. J. (2003, May 29–31). *Autonomous adaptive agent with grounded functional language*. Proceedings of the Seventh International Conference on Cognitive and Neural Systems, p. 43.

Hutchison, W. R., & Constantine, B. J. (2005). Selection by value: Biologically inspired learning for mobile robot control. *Journal of Intelligence Community Research and Development,* permanently available on Intelink.

Hutchison, W. R., Constantine, B. J., Borenstein, J., & Pratt, J. (2007, June). *Development of control for a serpentine robot*. Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2007), Jacksonville, FL.

Hutchison, W. R., Constantine, B. J., Borenstein, J., & Pratt, J. (2008a). Approach to validating a simulation of robot behavior: Human control of real robot replicates simulated robot behavior. *Journal of Intelligence Community Research and Development,* permanently available on Intelink.

Hutchison, W. R., Constantine, B. J., Borenstein, J., & Pratt, J. (2008b, November 10–12). *Developing control of a high-dof robot using reinforcement learning, genetic algorithms, scripting, and simulation*. Proceedings of the IEEE International Conference on Technologies for Practical Robot Applications, Woburn, MA.

Hutchison, W. R., & Stephens, K. R. (1987). Integration of distributed and symbolic knowledge representations. *Proceedings of the First International Conference on Neural Networks*, *2*, 395–398.

Levy, D. (2007). *Love and sex with robots: The evolution of human-robot relationships*. New York: HarperCollins.

Lin, L. (1992). Self-improving reactive agents based on reinforcement learning, planning, and teaching. *Machine Learning*, *8*, 293–321.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, *38*(8), 4–7.

Rachlin, H. (2012). Making IBM's computer, Watson, human. *The Behavior Analyst*, *35*, 1–16.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. R. Prokasy (Eds.), *Classical conditioning: II. Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Stephens, K. R., & Hutchison, W. R. (1992). Behavioral personal digital assistants: The seventh generation of computing. *The Analysis of Verbal Behavior*, *10*, 149–156.

Stephens, K. R., & Hutchison, W. R. (1993, October). Behavior analysis and the quest for machine intelligence. *Educational Technology*, pp. 52–61.

Stephens, K. R., Hutchison, W. R., Hormby, S., & Bell, T. M. (1990). Dynamic resource allocation using adaptive networks. *Neurocomputing*, *2*, 9–16.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Touretzky, D. S., & Saksida, L. M. (1997). Operant conditioning in Skinnerbots. *Adaptive Behavior*, *5*(3/4), 219–247.

Weng, J., & Chen, S. (1996). *Incremental learning for vision-based navigation* (Vol. 4, pp. 45–49). Proceedings of the International Conference on Pattern Recognition, Vienna, Austria.